

# Determinantal processes and completeness of random exponentials: the critical case

Subhroshekhar Ghosh  
 Department of Mathematics  
 University of California, Berkeley  
 subhro@math.berkeley.edu

## Abstract

For a locally finite point set  $\Lambda \subset \mathbb{R}$ , consider the collection of exponential functions given by  $\mathcal{E}_\Lambda := \{e^{i\lambda x} : \lambda \in \Lambda\}$ . We examine the question whether  $\mathcal{E}_\Lambda$  spans the Hilbert space  $L^2[-\pi, \pi]$ , when  $\Lambda$  is random. For several point processes of interest, this belongs to a certain critical case of the corresponding question for deterministic  $\Lambda$ , about which little is known. For  $\Lambda$  the continuum sine kernel process, obtained as the bulk limit of GUE eigenvalues, we establish that  $\mathcal{E}_\Lambda$  is indeed complete. We also answer an analogous question on  $\mathbb{C}$  for the Ginibre ensemble, arising as weak limits of certain non-Hermitian random matrix eigenvalues. In fact we establish completeness for any “rigid” determinantal point process in Euclidean space. Additionally, we answer two questions to Lyons and Steif about stationary determinantal processes on  $\mathbb{Z}^d$ .

# 1 Introduction

Any locally finite point set  $\Lambda \subset \mathbb{R}$  gives us a set of functions

$$\mathcal{E}_\Lambda := \{e_\lambda : \lambda \in \Lambda\} \subset L^2[-\pi, \pi],$$

where  $e_\lambda(x) = e^{i\lambda x}$ ,  $i$  being the imaginary unit. The following question is classical:

**Question 1.** *Does  $\mathcal{E}_\Lambda$  span  $L^2[-\pi, \pi]$  ?*

An equivalent term found in the literature to describe the fact that  $\mathcal{E}_\Lambda$  spans  $L^2[-\pi, \pi]$  is that  $\mathcal{E}_\Lambda$  is complete in  $L^2[-\pi, \pi]$ . When  $\Lambda$  is deterministic, this is a well studied problem in the literature. In the case where  $\Lambda$  is random, that is,  $\Lambda$  is a point process, much less is known. For any ergodic point process  $\Lambda$ , it can be easily checked that the event in question has a 0-1 law.

In this paper we provide a complete answer to Question 1 in the case where  $\Lambda$  is the continuum sine kernel process (see Section 2 for a precise definition):

**Theorem 1.1.** *When  $\Lambda$  is a realisation of the continuum sine kernel process on  $\mathbb{R}$ , a.s.  $\mathcal{E}_\Lambda$  spans  $L^2[-\pi, \pi]$ .*

Similar questions can be asked in higher dimensions as well. On  $\mathbb{C}$ , we consider the analogous question with  $\Lambda$  coming from the Ginibre ensemble (see Section 2 for a precise definition). An exponential function here is defined as  $\mathcal{E}_\lambda(z) = e^{\bar{\lambda}z}$  and the natural space to study completeness is the Fock-Bargmann space. The latter space is the closure of the set of polynomials (in one complex variable) in  $L^2(\gamma)$  where  $\gamma$  is the standard complex Gaussian measure on  $\mathbb{C}$ , having the density  $\frac{1}{\pi}e^{-|z|^2}$  with respect to the Lebesgue measure. Here we prove:

**Theorem 1.2.** *When  $\Lambda$  is a realisation of the Ginibre ensemble on  $\mathbb{C}$ , a.s.  $\mathcal{E}_\Lambda$  spans the Fock-Bargmann space. Equivalently, a.s. in  $\Lambda$  the following happens: if there is a function  $f$  in the Fock-Bargmann space which vanishes at all the points of  $\Lambda$ , then  $f \equiv 0$ .*

All these questions are specific realisations of the following completeness question that was asked of any determinantal process by Peres and Lyons in 2009.

Consider a determinantal point process  $\pi$  in a space  $\Xi$  equipped with a background measure  $\mu$ . Let  $\pi$  correspond to a projection onto the subspace  $\mathcal{H}$  of  $L^2(\mu)$  in the usual way, for details, see Section 2 and also [HKPV10], [Sos00] and [Ly03]. Let  $K(\cdot, \cdot)$  be the kernel of the determinantal process, which is also the integral kernel corresponding to the projection onto  $\mathcal{H}$ . Consequently,  $\mathcal{H}$  is a reproducing kernel Hilbert space, with the kernel  $K(\cdot, \cdot)$ . Let  $\{x_i\}_{i=1}^\infty$  be a sample from  $\pi$ . Clearly,  $\{K(x_i, \cdot)\}_{i=1}^\infty \subset \mathcal{H}$ . In 2009, Lyons and Peres asked the following question:

**Question 2.** *Is the random set of functions  $\{K(x_i, \cdot)\}_{i=1}^n$  complete in  $L^2(\mu)$  a.s.?*

The answer to Question 2 is trivial in the case where  $\mathcal{H}$  is finite dimensional (say dimension is  $N$ ), there it follows simply from the fact that the matrix  $(K(x_i, x_j))_{i,j=1}^N$  is a.s. non-singular. In the case where  $\Xi$  is a countable space, this was first proved for spanning forests by Morris [Mo03]. Subsequently, this has been answered in the affirmative for any discrete determinantal process by Lyons, see [Ly03]. However, in the continuum (e.g. when  $\Xi = \mathbb{R}^d$  and  $\mathcal{H}$  is infinite dimensional), the answer to Question 2 is unknown.

In this paper, we answer Question 2 in the affirmative for rigid determinantal processes. Let  $\pi$  be a translation invariant determinantal point process with determinantal kernel  $K(\cdot, \cdot)$  on  $\mathbb{R}^d$  with a background measure  $\mu$  that is mutually absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}^d$ , such that  $K(\cdot, \cdot)$  is also the projection kernel for the subspace  $\mathcal{H}$  that is canonically associated with  $\pi$ . Let the map  $x \rightarrow K(x, \cdot)$  be continuous as a map from  $\mathbb{R}^d \rightarrow \mathcal{H}$ .

**Theorem 1.3.** *Let  $\Pi$  be rigid, in the sense that for any ball  $B$ , the point configuration outside  $B$  a.s. determines the number of points  $N_B$  of  $\pi$  inside  $B$ . Then  $\{K(x, \cdot) : x \in \Pi\}$  is a.s. complete in  $\mathcal{H}$ .*

To see the correspondence between Question 2 and Theorems 1.1 - 1.2, we can make appropriate substitutions for the kernels and spaces in Theorem 1.3, for details see Section 2. The statement of Theorem 1.1 is equivalent to Question 2 (with an affirmative answer) under Fourier conjugation. The statement in Theorem 1.2 involving the vanishing of functions on  $\Lambda$  is a result of the fact that the Fock-Bargmann space is a reproducing kernel Hilbert space, with the reproducing kernel and background measure being the same as the determinantal kernel and the background measure of the Ginibre ensemble.

Completeness (in the appropriate Hilbert space) of collections of exponential functions indexed by a point configuration is a well-studied theme, for a classic reference see the survey by Redheffer [Re77]. However, most of the classical results deal with deterministic point configurations, and are often stated in terms of some sort of density of the underlying point set. E.g., one crucial parameter is the Beurling Malliavin density of the point configuration, for details, see [Ly03] Definition 7.13 and the ensuing discussion there. Typically, the results are of the following form : if the relevant density parameter is supercritical, then the exponential system is complete, and if it is subcritical, it is incomplete. E.g., see Beurling and Malliavin's theorem, stated as Theorem 71 in [Re77].

However, it is not hard to check that the point configurations of our interest almost surely exhibit the critical density in terms of the classical results. The critical density, in most cases, turns out to be equal to the one-point intensity for negatively associated ergodic point processes (including the homogeneous Poisson process). In this paper we have chosen the normalizations for our models such that the one-point intensity (and hence the critical density) is equal to 1. The critical cases in the deterministic setting are much harder to handle. E.g., in  $L^2[-\pi, \pi]$ ,  $\{e_\lambda : \lambda \in \mathbb{Z}\}$  is a complete set of exponentials, but  $\{e_\lambda : \lambda \in \mathbb{Z} \setminus \{0\}\}$  is incomplete. The study of completeness problems for random point configurations is much more limited, let alone in the critical case. Nevertheless, when the densities are super or subcritical, we can either refer to the results in the deterministic setting (e.g., Theorem 71 in [Re77]), or there is existing literature (see [CLP01] Theorems 1.1 and 1.2 or [SU97]). However, at critical densities, which is the case we are interested in, much less is known. To the best of our knowledge, the only resolved case is that of a

perturbed lattice, where completeness was established under some regularity conditions on the (random) perturbations, see [CL97] Theorem 5. Our result in Theorem 1.3 answers this question for natural point processes, like the Ginibre or the sine kernel, which are not i.i.d. perturbations of  $\mathbb{Z}$ .

There are other natural examples of determinantal point processes for which similar questions are not amenable to our approach. E.g., one can consider Question 2 for the zero process of the hyperbolic Gaussian analytic function, where the answer, either way, is unknown, because this determinantal point process is not rigid (see [HS10]).

En route proving Theorem 1.3, we provide a quantitative expression for the negative association property for determinantal point processes in the continuum, which is relevant for our purposes. In the discrete setting, a complete theorem to this effect has been proved in [Ly03]. However, we could not locate such a result in the literature on determinantal point processes in the continuum. In Theorem 1.4, we establish negative association for the number of points for determinantal point processes under minimal regularity hypotheses.

**Definition 1.** *We call two non-negative real valued random variables  $X$  and  $Y$  negatively associated if for any real numbers  $r$  and  $s$  we have*

$$\mathbb{P}((X \geq r) \cap (Y \geq s)) \leq \mathbb{P}(X \geq r)\mathbb{P}(Y \geq s).$$

*This is equivalent to the complementary condition*

$$\mathbb{P}((X \leq r) \cap (Y \leq s)) \leq \mathbb{P}(X \leq r)\mathbb{P}(Y \leq s).$$

In the theory of determinantal point processes on  $\mathbb{R}^d$ , by a standard kernel we mean a Hermitian kernel  $\mathbb{K}(x, y)$  which is continuous as a function from  $\mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{C}$  and is a non-negative trace class contraction when viewed as an integral operator from  $L^2(\mu) \rightarrow L^2(\mu)$  where  $\mu$  is the background measure. For further details, we refer the interested reader to [HKPV10].

**Theorem 1.4.** *For any determinantal point process with a standard kernel on  $\mathbb{R}^d$  and the background measure  $\mu$  absolutely continuous with respect to Lebesgue measure, the numbers of points in two disjoint Borel sets are negatively associated random variables, as in Definition 1.*

In addition to the completeness questions for random exponentials defined with respect to natural determinantal processes, we also answer two questions asked by Lyons and Steif in [LySt03]. First, we give a little background on a class of stationary determinantal processes on  $\mathbb{Z}^d$ .

Let  $f$  be a function  $\mathbb{T}^d \rightarrow [0, 1]$ . Then multiplication by  $f$  is a non-negative contraction operator from  $L^2(\mathbb{T}^d)$  to itself. Under Fourier conjugation, this gives rise to a non-negative contraction operator  $Q : \ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)$ . This, in turn, gives rise to a determinantal point process  $\mathbb{P}^f$  on  $\mathbb{Z}^d$  in a canonical way, for details see [Ly03]. For a point configuration  $\omega$  drawn from the distribution of  $\mathbb{P}^f$ , we denote by  $\omega(\underline{k})$  the indicator function of having a point at  $\underline{k} \in \mathbb{Z}^d$  in the configuration  $\omega$ . We denote by  $\omega_{\text{out}}$  the configuration of points on  $\mathbb{Z}^d \setminus \mathbf{0}$  obtained by restricting  $\omega$  to  $\mathbb{Z}^d \setminus \mathbf{0}$ , where  $\mathbf{0}$  denotes the origin. For a point

process  $\pi$  on a space  $\Xi$ , we denote by  $[\pi]$  the (random) counting measure obtained from a realisation of  $\pi$ . The  $k$ -point intensity functions of a point process, when it exists, will be denoted by  $\rho_k, k \geq 1$ . For the processes  $\mathbb{P}^f$ , all intensity functions exist. Moreover, translation invariance of  $\mathbb{P}^f$  implies that  $\rho_k(x_1 + x, \dots, x_k + x) = \rho_k(x_1, \dots, x_k)$  for all  $x, x_1, \dots, x_k \in \mathbb{Z}^d$ , in particular,  $\rho_1$  is a constant in  $[0, 1]$ .

In the paper [LySt03] it was conjectured that all determinantal processes obtained in this way are insertion and deletion tolerant, meaning that  $\mathbb{P}[\omega(\mathbf{0}) = 1 | \omega_{\text{out}}] > 0$  and  $\mathbb{P}[\omega(\mathbf{0}) = 0 | \omega_{\text{out}}] > 0$ . We answer this question in the negative, showing that for  $f$  which is the indicator function of an interval, this is not true.

**Theorem 1.5.** *Let  $f$  be the indicator function of an interval  $I \subset \mathbb{T}$ . Then there exists a measurable function*

$$N : \text{Point configurations on } \mathbb{Z} \setminus \mathbf{0} \rightarrow \mathbb{N} \cup \{0\}$$

*such that a.s. we have  $\omega(\mathbf{0}) = N(\omega_{\text{out}})$ . Consequently, the events  $\{\mathbb{P}[\omega(\mathbf{0}) = 1 | \omega_{\text{out}}] = 0\}$  and  $\{\mathbb{P}[\omega(\mathbf{0}) = 0 | \omega_{\text{out}}] = 0\}$  both have positive probability (in  $\omega_{\text{out}}$ ).*

We end by answering another question from [LySt03], where we demonstrate that “almost all” functions  $f$  can be reconstructed from the distribution  $\mathbb{P}^f$ .

**Theorem 1.6.** *Define  $\mathcal{E}$  to be the set of functions*

$$\mathcal{E} := \{f \in L^\infty(\mathbb{T}) : 0 \leq f(x) \leq 1 \text{ for all } x \in \mathbb{T}\}.$$

*Then  $\mathbb{P}^f$  determines  $f$  up to translation and flip, except possibly for a meagre set of functions in the  $L^\infty$  topology on  $\mathcal{E}$ .*

For any (as opposed to “almost all”) function  $f$ , we prove that  $\mathbb{P}^f$  determines the value distribution of  $f$ .

**Proposition 1.7.** *For any  $f \in \mathcal{E}$ ,  $\mathbb{P}^f$  determines the value distribution of  $f$ . This is true for  $\mathbb{P}^f$  defined on  $\mathbb{Z}^d$  for any  $d \geq 1$ .*

## 2 Definitions

In this section, we give precise descriptions of the models under study.

A determinantal point process on a space  $\Xi$  with background measure  $\mu$  and kernel  $\mathbb{K} : \Xi \times \Xi \rightarrow \mathbb{C}$ , is a point process whose  $n$ -point intensity functions (with respect to the measure  $\mu^{\otimes n}$ ) is given by

$$\rho_n(x_1, \dots, x_n) = \det \left( \mathbb{K}(x_i, x_j)_{i,j=1}^n \right).$$

The kernel  $\mathbb{K}$  induces an integral operator on  $L^2(\mu)$ , which must be locally trace class and, additionally, a positive contraction. An interesting class of examples is furnished when the integral operator given by  $\mathbb{K}$  is a projection onto a subspace  $\mathcal{H}$  of  $L^2(\mu)$ .

The Ginibre ensemble,  $K(z, w) = e^{\bar{z}w}$ ,  $d\mu(z) = e^{-|z|^2}d\mathcal{L}(z)$  and  $\mathcal{H}$  is the Fock-Bargmann space  $\subset L^2(\mu)$  (here  $\mathcal{L}$  is the Lebesgue measure on  $\mathbb{C}$ ). For every  $n$ , we can consider an  $n \times n$  matrix of i.i.d. complex Gaussian entries. The Ginibre ensemble arises as the weak limit (as  $n \rightarrow \infty$ ) of the point process given by the eigenvalues of this matrix. The continuum sine kernel process is given by  $K(x, y) = \frac{\sin \pi(x-y)}{\pi(x-y)}$ ,  $\mu$  = the Lebesgue measure on  $\mathbb{R}$ ,  $\mathcal{H}$  = the Fourier conjugates of the set of  $L^2$  functions supported on  $[-\pi, \pi]$ . The continuum sine kernel process arises as the bulk limit of the eigenvalues of the Gaussian Unitary Ensemble (GUE).

For greater details on these processes, see [HKPV10].

### 3 Completeness of random function spaces

In this Section, we prove Theorem 1.3. Due to the connections discussed in the introduction, this will automatically establish Theorems 1.1 and 1.2. On the way, we provide a proof of Theorem 1.4.

Before the main theorem, we establish a preparatory result.

The first of these results is a technical one:

**Proposition 3.1.** *Suppose  $\Pi$  is a rigid point process, in the sense that for any ball  $B$ , the point configuration outside  $B$  a.s. determines the number of points  $N_B$  of  $\Pi$  inside  $B$ . Assume that for any ball  $B$ ,  $\mathbb{E}[N_B] < \infty$ . Let  $A(r)$  denote the closed annulus of thickness  $r$  around  $B$ . Then we have*

$$\mathbb{E}[N_B | \Pi|_{A(r)}] \rightarrow N_B \quad (1)$$

a.s. as  $r \rightarrow \infty$ .

*Proof.* This follows from the convergence of the Doob's martingales  $M_r := \mathbb{E}[N_B | \Pi|_{A(r)}]$ ,  $r \geq 0$  of  $N_B$ , as  $r \rightarrow \infty$ . Note that  $M_\infty = N_B$  because  $\Pi$  is rigid. ■

*Proof of Theorem 1.4.* We will use an appropriate discretization argument to harness the result for the discrete case (see Theorem 8.1, [Ly03]), and pass to the continuum limit.

Let  $A, B$  be disjoint Borel sets in  $\mathbb{R}^d$  and  $r, s$  be two non-negative integers. We intend to prove

$$\mathbb{P}((N(A) \leq r) \cap (N(B) \leq s)) \leq \mathbb{P}(N(A) \leq r) \mathbb{P}(N(B) \leq s) \quad (2)$$

First of all, we can assume the sets  $A, B$  to be contained in a compact  $d$ -dimensional cube  $\mathfrak{D}$ , the general case can easily be deduced from this by considering  $A \cap \mathfrak{D}$  and  $B \cap \mathfrak{D}$  and letting  $\mathfrak{D} \uparrow \mathbb{R}^d$ ; the probabilities in (2) are converge to the appropriate limits under this procedure.

On  $L^2(\mu|_{\mathfrak{D}})$  we have  $\mathbb{K}$  is again a standard kernel and  $\mu|_{\mathfrak{D}}$  is clearly absolutely continuous with respect to Lebesgue measure restricted to  $\mathfrak{D}$ . Therefore, by Mercer's theorem we have an eigenvector expansion for the kernel

$$\mathbb{K}(x, y) = \sum_{\lambda_i \downarrow 0} \lambda_i \phi_i(x) \overline{\phi_i(y)} \quad (3)$$

where  $\phi_i$  are the eigenvectors and  $\lambda_i$  the corresponding eigenvalues for  $\mathbb{K}$  acting from  $L^2(\mu|_{\mathfrak{D}})$  to itself. For brevity, from here on we will drop the subscript  $\mathfrak{D}$  and pretend that  $\mu$  is a measure supported on  $\mathfrak{D}$ .

Both sides of (2) are continuous in the kernel  $\mathbb{K}$ , in the sense that if  $\mathbb{K}_n$  is a sequence of standard kernels converging to  $\mathbb{K}$  as continuous functions on  $\mathfrak{D} \times \mathfrak{D}$ , then the corresponding probabilities in (2) converge to that of  $\mathbb{K}$ . This enables us to replace  $k$  by finite truncations of the expansion (3). Since  $\mathbb{K}$  is a standard kernel, we have  $0 < \lambda_i \leq 1$ . However, for technical reasons we will assume that  $\lambda_i < 1$  for all  $i$ . The general case can be easily recovered by taking limits  $\lambda_i \uparrow 1$  for the eigenvalue 1, this can again be justified by the regularity of (2) under  $\mathbb{K}$ .

We have thus reduced to the case when  $\mathbb{K}(x, y) = \sum_{i=1}^N \lambda_i \phi_i(x) \overline{\phi_i(y)}$  where  $N$  is a positive integer and  $0 < \lambda_i < 1$ , without loss of generality say  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ . Moreover, the background measure and the ambient  $L^2$  space lives on  $\mathfrak{D}$ .

For a positive integer  $m$ , we divide each side of  $\mathfrak{D}$  into  $m$  equal parts, and index the resulting sub-cubes by  $\mathfrak{D}_i; i = 1, 2, \dots, n = m^d$ . We define a measure  $\mu_n$  on  $[n]$  by setting  $\mu_n(i) = \mu(\mathfrak{D}_i)$ . If the centre of  $\mathfrak{D}_i$  is  $x_i$ , then we define a function  $\mathbb{K}_n : [n] \times [n] \rightarrow \mathbb{C}$  by  $\mathbb{K}_n(i, j) = \mathbb{K}(x_i, x_j) \sqrt{\mu_n(i) \mu_n(j)}$ .

We claim that for  $n$  large enough,  $\mathbb{K}_n$  as defined above is (almost) a positive contraction on  $\ell_n^2$ . To see this, consider the quadratic form  $Q(a_1, \dots, a_n) = \sum_{i,j=1}^n a_i \overline{a_j} \mathbb{K}_n(i, j)$ . Clearly, the matrix  $\mathbb{K}_n$  is Hermitian, and we intend to show that  $|Q(a_1, \dots, a_n)| \leq 1$  when  $\|(a_1, \dots, a_n)\|_2 \leq 1$ . Observe that the  $a_i$  for which  $\mu_n(i) = 0$  do not matter in  $Q$ , so we can scale  $a_i$  by  $\sqrt{\mu_n(i)}$  and consider the form  $Q'(a'_1, \dots, a'_n) = \sum_{i,j=1}^n a'_i \overline{a'_j} \mathbb{K}_n(i, j) \sqrt{\mu_n(i) \mu_n(j)}$ . We want to maximise  $|Q'(a'_1, \dots, a'_n)|$  over  $\sum_{i=1}^n |a'_i|^2 \mu_n(i) \leq 1$ . We compare  $Q'(a'_1, \dots, a'_n)$  with the integral  $I(a') = \int \int_{\mathfrak{D} \times \mathfrak{D}} a'(x) \mathbb{K}(x, y) \overline{a'(y)} d\mu(x) d\mu(y)$ , where  $a'(x) = \sum_{i=1}^n a'_i \chi_{\mathfrak{D}_i}(x)$ . Here  $\chi_{\mathfrak{D}_i}$  is, as usual, the indicator function of  $\mathfrak{D}_i$ . For any  $\delta > 0$ , by taking  $n$  large enough, we can ensure that  $\sup_{(x,y) \in \mathfrak{D}_i \times \mathfrak{D}_j} |\mathbb{K}(x, y) - \mathbb{K}(x_i, x_j)| < \delta$ , hence

$$|I(a') - Q'(a'_1, \dots, a'_n)| \leq \delta \left( \sum_{i=1}^n |a'_i| \mu_n(i) \right)^2 \leq \delta \left( \sum_{i=1}^n |a'_i|^2 \mu_n(i) \right) \left( \sum_{i=1}^n \mu_n(i) \right) \leq \delta \mu(\mathfrak{D})$$

Further, note that since the spectral norm of  $\mathbb{K}$  as an integral operator on  $L^2(\mu \times \mu)$  is  $\lambda_1 < 1$ , therefore,  $|I(a)| \leq \lambda_1 < 1$ . Hence, for large enough  $n$ , we can have  $\delta$  small enough such that

$$|Q'(a'_1, \dots, a'_n)| \leq I(a') + \delta \mu(\mathfrak{D}) \leq \lambda_1 + \delta \mu(\mathfrak{D}) \leq 1$$

Hence for  $n$  large enough,  $\mathbb{K}_n$  is a contraction. A similar argument would show that  $\inf_{\|(a_1, \dots, a_n)\|_2=1} Q(a_1, \dots, a_n) > -\delta \mu(\mathfrak{D})$ . Because,  $\inf I(f) > 0$  over  $f \in L^2(\mu)$  with  $\|f\|_2 = 1$  is  $\geq 0$ . Hence,  $\mathbb{K}'_n = \mathbb{K}_n + \delta \mu(\mathfrak{D}) I_n$  is a positive contraction, where  $I_n$  is the identity matrix of size  $n$ .

Thus, we can consider a discrete determinantal process  $\mathbb{P}_n$  on  $\ell^2([n])$  given by the contraction  $\mathbb{K}'_n$  (for a fixed  $\delta$ ) and from Theorem 8.1 [Ly03] we know that such a process satisfies

$$\mathbb{P}_n((N(A) \leq r) \cap (N(B) \leq s)) \leq \mathbb{P}_n(N(A) \leq r) \mathbb{P}(N(B) \leq s) \quad (4)$$

where  $A$  and  $B$  are any two disjoint subsets  $[n]$  and  $N(A)$  is the number of indices  $\in A$  obtained in a random sample from  $\mathbb{P}_n$ .

To obtain the final result, we fix two disjoint subsets  $A$  and  $B$  of  $\mathfrak{D}$  which are finite unions of dyadic subcubes of  $\mathfrak{D}$ . It suffices to prove (2) for such sets because of the regularity of both sides of (2) under approximations  $A_n \uparrow A$ . Now, for large enough dyadic partition of size  $n$  (such that all the dyadic sub cubes comprising  $A$  and  $B$  are present in this partition), we can consider the corresponding determinantal process  $\mathbb{P}_n$ . Denoting by  $A$  and  $B$  the set of indices corresponding to the sub cubes of the sets  $A$  and  $B$  respectively, we can write down (4) where  $N(A)$  is as usual the number of indices in  $A$  obtained in a random sample from  $\mathbb{P}_n$ . Now, in the continuum, the probabilities  $\mathbb{P}((N(A) \leq r) \cap (N(B) \leq s))$  and  $\mathbb{P}(N(A) \leq r)$  can be expressed as certain integrals of  $\mathbb{K}$  whereas for  $\mathbb{P}_n$  the corresponding probabilities are given by certain finite sums. As we let first  $n \rightarrow \infty$  and then  $\delta \rightarrow 0$ , the sums converge to the corresponding integrals. So, letting  $n \rightarrow \infty$  (along powers of 2) in (4), we obtain (2) in the limit.

This completes the proof of the theorem. ■

We are now ready to establish the main Theorem 1.3.

*Proof of Theorem 1.3.* Let  $\mathcal{H}_0$  be the random closed subspace of  $\mathcal{H}$  generated by the functions  $\{K(x, \cdot) : x \in \Pi\}$  inside  $L^2(\mu)$ . We wish to show that a.s. we have  $\mathcal{H}_0 = \mathcal{H}$ . It suffices to prove that  $K(\mathbf{0}, \cdot) \in \mathcal{H}_0$  a.s., where  $\mathbf{0}$  denotes the origin in  $\mathbb{R}^d$ . Because, by translation invariance, this would imply that  $K(q, \cdot) \in \mathcal{H}_0$  a.s. for all  $q \in \mathbb{R}^d$  with all co-ordinates rational. Then the continuity of the map  $x \rightarrow K(x, \cdot)$  would imply that  $K(x, \cdot) \in \mathcal{H}_0$  simultaneously for all  $x \in \mathbb{R}^d$ . This implies that  $\mathcal{H}_0 = \mathcal{H}$  a.s.

For the points  $\{x_i\}_{i=1}^n \in \mathbb{R}^d$ , let  $\mathfrak{D}(x_1, \dots, x_n)$  denote  $\text{Det} \left[ (K(x_i, x_j))_{i,j=1}^n \right]$ . Then, if we consider the functions  $K(z, \cdot)$  as a vector in the Hilbert space  $\mathcal{H}$ , the squared norm of the projection of  $K(x, \cdot)$  on to the orthogonal complement of  $\text{Span} \{K(x_i, \cdot), 1 \leq i \leq n\}$  is given by the ratio  $\frac{\mathfrak{D}(x, x_1, \dots, x_n)}{\mathfrak{D}(x_1, \dots, x_n)}$ . But this is also equal to the conditional intensity  $p(x|x_1, \dots, x_n)$  of  $\pi$  at  $x$  given that  $\{x_1, \dots, x_n\} \subset \Pi$ .

Fix an  $\epsilon > 0$ . For  $r > 0$ , denote by  $B_r$  the ball of radius  $r$  centred at  $\mathbf{0}$ . Let  $\omega_R$  be the set of points of  $\Pi$  in  $B_R \setminus B_\epsilon$ . For any feasible realisation  $\Upsilon_R$  of  $\omega_R$  we have

$$\mathbb{E}[\text{Number of points in } B_\epsilon | \Upsilon_R \subset \Pi] = \int_{B_\epsilon} p(x | \Upsilon_R) d\mu(x) \quad (5)$$

We now proceed with the left hand side in (5) as

$$\begin{aligned}
& \mathbb{E}[\text{Number of points in } B_\epsilon | \Upsilon_R \subset \Pi] \\
&= \sum_{k=1}^{\infty} \mathbb{P}[\text{There are } \geq k \text{ points in } B_\epsilon | \Upsilon_R \subset \Pi] \\
&\leq \sum_{k=1}^{\infty} \mathbb{P}[\text{There are } \geq k \text{ points in } B_\epsilon | \omega_R = \Upsilon_R] \text{ (from negative association, see Theorem 1.4)} \\
&= \mathbb{E}[\text{Number of points in } B_\epsilon | \omega_R = \Upsilon_R]
\end{aligned}$$

The inequality above involving negative association follows by applying Theorem 1.4 to the point process obtained by conditioning  $\Pi$  to contain  $\Upsilon_R$ , which is a determinantal process on  $\mathbb{R}^d$  having a standard kernel, see [ShTa03] Corollary 6.6. By Proposition 3.1, we have given any  $\delta > 0$ , we can find  $R_\delta$  such that except on an event  $\Omega_1^\delta$  of probability  $< \delta$ , we have

$$|\mathbb{E}[\text{Number of points in } B_\epsilon | \omega_{R_\delta}] - N_{B_\epsilon}| < \delta.$$

But, except on an event  $\Omega_2$ , measurable with respect to  $\omega_{\text{out}}$  and of probability  $O(\epsilon^d)$ , we have  $N_{B_\epsilon} = 0$ .

Hence, except on the event  $\Omega_1^\delta \cup \Omega_2$ , we have

$$\int_{B_\epsilon} p(x | \Upsilon_{R_\delta}) d\mu(x) \leq \delta$$

Letting  $\delta \downarrow 0$  along a summable sequence, we deduce that a.s. on  $(\Omega_2)^c$  we have

$$\varliminf_{\delta \rightarrow 0} \int_{B_\epsilon} p(x | \Upsilon_{R_\delta}) d\mu(x) = 0.$$

By Fatou's lemma, this implies that on  $(\Omega_2)^c$  we have

$$\int_{B_\epsilon} \varliminf_{\delta \rightarrow 0} p(x | \Upsilon_{R_\delta}) d\mu(x) = 0.$$

This implies that  $\varliminf_{\delta \rightarrow 0} p(x | \Upsilon_{R_\delta}) = 0$  for almost every  $x \in B_\epsilon$  on the event  $(\Omega_2)^c$ . By our previous discussion on the connection between the squared norms of projections and conditional intensities, this means that on  $\Omega_2^c$ ,  $K(x, \cdot) \in \mathcal{H}_0$  for a dense set of  $x \in B_\epsilon$ . By the continuity of the map  $x \rightarrow K(x, \cdot)$ , we have  $K(\mathbf{0}, \cdot) \in \mathcal{H}_0$  on the event  $\Omega_2^c$ . Letting  $\epsilon \rightarrow 0$  (which implies  $\Omega_2 \downarrow \phi$ ), we have  $K(\mathbf{0}, \cdot) \in \mathcal{H}_0$  with probability one.  $\blacksquare$

## 4 Rigidity and Tolerance for certain determinantal point processes

In this Section we discuss the insertion and deletion tolerance question from [LySt03], principally the proof of Theorem 1.5.

We will use the following general observation for determinantal point processes given by a projection kernel:

**Proposition 4.1.** *Consider a determinantal point process  $\Pi$  in a locally compact space  $\Xi$  with determinantal kernel  $K(\cdot, \cdot)$  and background measure  $\mu$ , such that  $K$  is idempotent as an integral operator on  $L^2(\mu)$ . Let  $\psi$  be a compactly supported function on  $\Xi$ . Then*

$$\text{Var} \left[ \int \psi d[\Pi] \right] = \frac{1}{2} \iint |\psi(x) - \psi(y)|^2 |K(x, y)|^2 d\mu(x) d\mu(y) \quad (6)$$

*Proof.* This follows from the determinantal formula for the two point intensity function of  $\Pi$  and the idempotence of  $K$ . ■

*Proof of Theorem 1.5.* We will approach this question by estimating the variance of linear statistics of  $\mathbb{P}^f$ . A similar approach has been used in [GP12] to obtain rigidity behaviour for the Ginibre ensemble and the zero process of the standard planar Gaussian analytic function.

Let  $\varphi$  be a  $C_c^\infty$  function on  $\mathbb{R}$  which is  $\equiv 1$  in a neighbourhood of the origin. Viewed as a function on  $\mathbb{Z}$ ,  $\varphi$  is compactly supported and  $= 1$  at the origin. Let  $\varphi_L$  be defined by  $\varphi_L(x) = \varphi(x/L)$ . Since  $f$  is the indicator function of an interval  $I \subset \mathbb{T}$ , therefore the determinantal kernel  $K$  of  $\mathbb{P}^f$  is an idempotent on  $\ell^2(\mathbb{Z})$ . Applying Proposition 4.1 with  $\Pi = \mathbb{P}^f$ ,  $\Xi = \mathbb{Z}$ ,  $\mu =$  the counting measure on  $\mathbb{Z}$  and  $\psi = \varphi_L$  we get

$$\text{Var} \left[ \int \varphi_L d[\mathbb{P}^f] \right] = \frac{1}{2} \sum_{i, j \in \mathbb{Z}} |\varphi(\frac{i}{L}) - \varphi(\frac{j}{L})|^2 |\hat{f}(i - j)|^2. \quad (7)$$

Observe that rotating the interval  $I \subset \mathbb{T}$  leaves the measure  $\mathbb{P}^f$  invariant, so, without loss of generality, we take  $I$  to be corresponding to the interval  $[-a, a]$  where  $\mathbb{T}$  is parametrized as  $(-\pi, \pi]$  and  $0 < a < \pi$ . Then  $\hat{f}(k) = c(a) \sin ak/k$  where  $c(a)$  is a constant. This implies that, for some constant  $c > 0$ , we have

$$\text{Var} \left[ \int \varphi_L d[\mathbb{P}^f] \right] = c \sum_{i, j \in \mathbb{Z}} |\varphi(\frac{i}{L}) - \varphi(\frac{j}{L})|^2 (\sin^2 a(i - j)) |i - j|^{-2}. \quad (8)$$

This, in turn, implies

$$\text{Var} \left[ \int \varphi_L d[\mathbb{P}^f] \right] \leq c \sum_{i, j \in \mathbb{Z}} |\varphi(\frac{i}{L}) - \varphi(\frac{j}{L})|^2 |(i - j)/L|^{-2} L^{-2}. \quad (9)$$

Hence we have

$$\overline{\lim}_{L \rightarrow \infty} \text{Var} \left[ \int \varphi_L d[\mathbb{P}^f] \right] \leq c \int \int \left( \frac{\varphi(x) - \varphi(y)}{x - y} \right)^2 d\mathcal{L}(x) d\mathcal{L}(y) \quad (10)$$

where  $\mathcal{L}$  denotes the Lebesgue measure on  $\mathbb{R}$ .

For any  $\mathbb{C}_c^1$  functions  $\psi_1, \psi_2$  on  $\mathbb{R}$ , we define the form

$$\Lambda(\psi_1, \psi_2) = \int \int \frac{(\psi_1(x) - \psi_1(y))(\psi_2(x) - \psi_2(y))}{(x - y)^2} d\mathcal{L}(x) d\mathcal{L}(y). \quad (11)$$

It is known that  $\Lambda(\psi, \psi)$  is related to the  $H^{1/2}$  norm of  $\psi$ .

A simple calculation shows that for any  $\lambda > 0$ , we have  $\Lambda((\psi_1)_\lambda, (\psi_2)_\lambda) = \Lambda(\psi_1, \psi_2)$ . In particular, this implies that  $\Lambda(\psi_1, (\psi_2)_\lambda) = \Lambda((\psi_1)_{1/\lambda}, \psi_2)$ . Further, we will see in Proposition 4.3 that  $\Lambda(\psi, \psi_\lambda) \rightarrow 0$  as  $\lambda \rightarrow 0$ .

For an integer  $n > 0$ , let  $0 < \lambda = \lambda(n) < 1$  be such that for  $\varphi$  as above,  $|\Lambda(\varphi, \varphi_{\lambda^{-i}})| \leq 1/2^i$  for  $1 \leq i \leq n$ . Such a choice can be made because of the observations in the previous paragraph. Define  $\Phi^n = (\sum_{i=1}^n \varphi_{\lambda^{-i}})/n$ . Note that  $\Phi^n \equiv 1$  in a neighbourhood of  $\mathbf{0}$  in  $\mathbb{R}$ , and the same is true for all scalings  $\Phi_L^n$  of  $\Phi^n$  whenever  $L \geq 1$ .

Let  $L > 1$  be such that

$$\text{Var} \left[ \int \Phi_L^n d[\mathbb{P}^f] \right] \leq c\Lambda(\Phi^n, \Phi^n) + \frac{1}{n}.$$

But  $\Lambda(\Phi^n, \Phi^n) = \frac{1}{n^2} \left( \sum_{i,j=1}^n \Lambda(\varphi_{\lambda^{-i}}, \varphi_{\lambda^{-j}}) \right)$ . Observe that

$$\Lambda(\varphi_{\lambda^{-i}}, \varphi_{\lambda^{-j}}) = \Lambda(\varphi, \varphi_{\lambda^{-|i-j|}}) \leq 2^{-|i-j|}.$$

This implies that

$$\sum_{i,j=1}^n \Lambda(\varphi_{\lambda^{-i}}, \varphi_{\lambda^{-j}}) \leq C(\varphi)n.$$

Hence  $\text{Var} \left[ \int \Phi_L^n d[\mathbb{P}^f] \right] \leq C(\varphi)/n$ .

By the Borel-Cantelli lemma, we have, as  $n \rightarrow \infty$ ,

$$\left| \int \Phi_L^{2^n} d[\mathbb{P}^f] - \mathbb{E} \left[ \int \Phi_L^{2^n} d[\mathbb{P}^f] \right] \right| \rightarrow 0. \quad (12)$$

But  $\int \Phi_L^{2^n} d[\mathbb{P}^f] = \omega(\mathbf{0}) + \int_{\mathbb{Z} \setminus \mathbf{0}} \Phi_L^{2^n} d[\mathbb{P}^f]$ , and the second term can be evaluated if we know  $\omega_{\text{out}}$ .  $\mathbb{E} \left[ \int \Phi_L^{2^n} d[\mathbb{P}^f] \right]$  can also be computed explicitly in terms of the first intensity measure of  $\mathbb{P}^f$ . This implies that from (12), we can deduce the value of  $\omega(\mathbf{0})$  by letting  $n \rightarrow \infty$ .

Thus,  $\omega_{\text{out}}$  a.s. determines the value of  $\omega(\mathbf{0})$ . Since both the events  $\omega(\mathbf{0}) = 0$  and  $\omega(\mathbf{0}) = 1$  occur with positive probability, therefore the events  $\{\mathbb{P}[\omega(\mathbf{0}) = 1 | \omega_{\text{out}}] = 0\}$  and  $\{\mathbb{P}[\omega(\mathbf{0}) = 1 | \omega_{\text{out}}] = 0\}$  both have positive probability (in  $\omega_{\text{out}}$ ).  $\blacksquare$

**Remark 4.1.** For  $f$  which is the indicator function of a finite, disjoint union of intervals  $\subset \mathbb{T}$ , we have  $|\hat{f}(k)| \leq c/|k|$ , hence the same argument and the same conclusion as Theorem 1.5 holds for such  $f$ .

**Remark 4.2.** A similar argument shows that, in fact, for any finite set  $S \subset \mathbb{Z}$ , the point configuration of  $\mathbb{P}^f$  restricted to  $S^c$  a.s. determines the number of points of  $\mathbb{P}^f$  in  $S$ , where  $f$  is the indicator function of an interval.

A similar class of determinantal point processes in the continuum can be obtained by considering  $L^2$  functions  $f : \mathbb{R}^d \rightarrow [0,1]$ . The multiplication operator  $M_f$  defined by such a function  $f$  is clearly a contraction on  $L^2(\mathbb{R}^d)$ . By considering the Fourier conjugate of such an operator, we get another contraction on  $L^2(\mathbb{R}^d)$ , which gives us a translation invariant

determinantal point process  $\mathbb{P}^f$  in  $\mathbb{R}^d$ . One of the most important examples of such a point process is the sine kernel process on  $\mathbb{R}$ , which is defined by the determinantal kernel  $\frac{\sin \pi(x-y)}{\pi(x-y)}$  with the Lebesgue measure on  $\mathbb{R}$  as the background measure. Here the relevant function  $f$  is the indicator function of the interval  $[-\pi, \pi]$ . For details, see [AGZ09]. More generally we can consider the indicator function of any measurable subset of  $\mathbb{R}$ , which will give us a projection operator on  $L^2(\mathbb{R})$ , and hence a determinantal point process corresponding to a projection kernel. In this setting, we have a continuum analogue of Theorem 1.5, which says that whenever  $f$  is the indicator of a finite union of compact intervals in  $\mathbb{R}$ , we have that  $\mathbb{P}^f$  is a rigid process.

**Theorem 4.2.** *Let  $f : \mathbb{R} \rightarrow [0, 1]$  be an indicator function of a finite union of compact intervals. Then the determinantal point process  $\mathbb{P}^f$  in  $\mathbb{R}$  is “rigid” in the following sense. Let  $U \subset \mathbb{R}$  be an interval, and let  $\omega$  be the point configuration sampled from the distribution  $\mathbb{P}^f$ . Define the restricted point configurations  $\omega_{\text{in}} = \omega|_U$  and  $\omega_{\text{out}} = \omega|_{U^c}$ . Let  $|\omega_{\text{in}}|$  be the number of points of  $\omega$  in  $U$ . Then there exists a measurable function*

$$N : \text{Point configurations in } U^c \rightarrow \mathbb{N} \cup \{0\}$$

*such that a.s. we have  $|\omega_{\text{in}}| = N(\omega_{\text{out}})$ . This holds true for all intervals  $U$ . In particular, the continuum sine kernel process is “rigid” in the above sense.*

*Proof.* By translation invariance, it suffices to take  $U$  to be centred at the origin. Let  $\varphi$  be a  $C_c^\infty$  function which is  $\equiv 1$  in a neighbourhood of  $U$ . Then we have

$$\text{Var} \left[ \int \varphi d[\mathbb{P}^f] \right] = \int \int |\varphi(x) - \varphi(y)|^2 |\hat{f}(x-y)|^2 d\mathcal{L}(x) d\mathcal{L}(y). \quad (13)$$

But for any compact interval  $[a, b]$  we have  $|\hat{1}_{[a,b]}(\xi)| \leq c|\xi|^{-1}$ , hence we have

$$\text{Var} \left[ \int \varphi d[\mathbb{P}^f] \right] \leq C \int \int \left( \frac{\varphi(x) - \varphi(y)}{x-y} \right)^2 d\mathcal{L}(x) d\mathcal{L}(y). \quad (14)$$

Recall the form  $\Lambda(\cdot, \cdot)$  as in (11). Proposition 4.3 implies that  $\Lambda(\varphi, \varphi_{\lambda^{-1}}) \rightarrow 0$  as  $\lambda \rightarrow 0$ . For an integer  $n > 0$ , let  $0 < \lambda < 1$  be such that for  $\varphi$  as above,  $|\Lambda(\varphi, \varphi_{\lambda^{-i}})| \leq 1/2^i$  for  $1 \leq i \leq n$ . Define  $\Phi^n = (\sum_{i=1}^n \varphi_{\lambda^{-i}})/n$ . We have  $\Phi^n \equiv 1$  in a neighbourhood of  $U$  in  $\mathbb{R}$ . Due to our choice of  $\lambda$ , we have  $\text{Var}(\int \Phi^n d[\mathbb{P}^f]) = \Lambda(\Phi^n, \Phi^n) = O(1/n)$ . From here, we proceed on similar lines to the proof of Theorem 1.5 and deduce the existence of  $N$  as prescribed in the statement of Theorem 4.2.  $\blacksquare$

We now prove Proposition 4.3, which will complete the proof of Theorem 1.5.

**Proposition 4.3.** *For a  $C_c^1$  function  $\varphi$ , we have  $\Lambda(\varphi, \varphi_{\lambda^{-1}}) \rightarrow 0$  as  $\lambda \rightarrow 0$ .*

*Proof.* We begin with the expression

$$\Lambda(\varphi, \varphi_{\lambda^{-1}}) = \int \int \frac{(\varphi(x) - \varphi(y))(\varphi(\lambda x) - \varphi(\lambda y))}{(x-y)^2} d\mathcal{L}(x) d\mathcal{L}(y). \quad (15)$$

Fix  $a > 0$ . Let  $K$  be the support of  $\varphi$ . We define the function

$$\gamma(x, y) = \begin{cases} \|\varphi'\|_\infty^2 & \text{if } x \text{ or } y \in K \text{ and } |x - y| \leq a \\ \frac{4\|\varphi\|_\infty^2}{(x-y)^2} & \text{if } x \text{ or } y \in K \text{ and } |x - y| > a \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

For  $0 < \lambda < 1$ , the integrand in (15) is bounded from above pointwise by  $\gamma(x, y)$ . To see this, note that the integrand in (15) is non-zero only on the set  $S = \{(x, y) : x \text{ or } y \in K\}$ . On  $S$ , we bound the integrand from above as follows: for  $(x, y) \in S$  such that  $|x - y| \leq a$  we use  $|\varphi(x) - \varphi(y)| \leq \|\varphi'\|_\infty |x - y|$ , for other  $(x, y) \in S$  we use  $|\varphi(x) - \varphi(y)| \leq 2\|\varphi\|_\infty$ .

Since  $K$  is a compact set, we have

$$\int \int \gamma(x, y) d\mathcal{L}(x) d\mathcal{L}(y) < \infty, \quad (17)$$

where we use the fact that  $\int_{|t|>a} \frac{1}{t^2} dt = \frac{1}{a}$ .

(17) enables us to use the dominated convergence theorem and let  $\lambda \rightarrow 0$  in the integrand of (15), whence  $\Lambda(\varphi, \varphi_{\lambda^{-1}}) \rightarrow 0$ .  $\blacksquare$

Next we show that in any dimension  $d$ , whenever  $f$  is not the indicator of a subset of  $\mathbb{T}^d$ ,  $\text{Var} [\int \varphi_L d[\mathbb{P}^f]]$  blows up at least like  $L^d$  as  $L \rightarrow \infty$ .

**Proposition 4.4.** *Let  $f : \mathbb{T}^d \rightarrow [0, 1]$  not equal the indicator function of some subset of  $\mathbb{T}^d$  (up to Lebesgue-null sets). Then  $\text{Var} [\int \varphi_L d[\mathbb{P}^f]] = \Omega(L^d)$  as  $L \rightarrow \infty$ .*

*Proof.* Let  $\rho_2(\cdot, \cdot)$  be the two point intensity function of  $\mathbb{P}^f$ , given by the formula

$$\rho_2(i, j) = \det \begin{pmatrix} \hat{f}(0) & \hat{f}(i - j) \\ \hat{f}(j - i) & \hat{f}(0) \end{pmatrix}$$

Let  $\lambda_d$  denote the normalized Lebesgue measure on  $\mathbb{T}^d$ . Using the above formula, we can write the variance in question as:

$$\begin{aligned} \text{Var} \left[ \int \varphi_L d[\mathbb{P}^f] \right] &= \mathbb{E} \left( \int \varphi_L d[\mathbb{P}^f] \right)^2 - \left( \mathbb{E} \left[ \int \varphi_L d[\mathbb{P}^f] \right] \right)^2 \\ &= \sum_i \varphi_L(i)^2 \hat{f}(0) + \sum_{i,j} \varphi_L(i) \varphi_L(j) \left( \hat{f}(0)^2 - |\hat{f}(i - j)|^2 \right) - \sum_{i,j} \varphi_L(i) \varphi_L(j) \hat{f}(0)^2 \\ &= \sum_i \varphi_L(i)^2 \left( \hat{f}(0) - \sum_j |\hat{f}(i - j)|^2 \right) + \sum_{i,j} (\varphi_L(i)^2 - \varphi_L(i) \varphi_L(j)) |\hat{f}(i - j)|^2 \\ &= \left( \hat{f}(0) - \sum_k |\hat{f}(k)|^2 \right) \left( \sum_k \varphi_L(k)^2 \right) + \frac{1}{2} \left( \sum_{i,j} |\varphi_L(i) - \varphi_L(j)|^2 |\hat{f}(i - j)|^2 \right) \\ &\geq \left( \hat{f}(0) - \sum_k |\hat{f}(k)|^2 \right) \left( \sum_k \varphi_L(k)^2 \right) \\ &= \left( \int_{\mathbb{T}^d} f(x) d\lambda_d(x) - \int_{\mathbb{T}^d} f(x)^2 d\lambda_d(x) \right) \left( \sum_k \varphi_L(k)^2 \right). \end{aligned}$$

In the last step we have used Parseval's identity:  $\sum_k |\hat{f}(k)|^2 = \int f(x)^2 d\lambda_d(x)$ . Note that since  $0 \leq f \leq 1$ , we have  $\left( \int_{\mathbb{T}^d} f(x) d\lambda_d(x) - \int_{\mathbb{T}^d} f(x)^2 d\lambda_d(x) \right) \geq 0$  with strict inequality holding if and only if  $f$  is not the indicator of some subset of  $\mathbb{T}^d$ . Finally, observe that as  $L \rightarrow \infty$  we have

$$\frac{1}{L^d} \left( \sum_k \varphi_L(k)^2 \right) = \sum_k \frac{1}{L^d} \varphi \left( \frac{k}{L} \right)^2 \rightarrow \|\varphi\|_2^2.$$

This completes the proof of the proposition. ■

## 5 $\mathbb{P}^f$ determines $f$

In this Section we provide the proofs of Theorem 1.6 and Proposition 1.7

*Proof of Theorem 1.6.* Define  $\mathfrak{F}$  to be the subset of functions of  $\mathcal{E}$  satisfying the following conditions:

- (i) Either  $\hat{f}(k) \neq 0 \forall k \in \mathbb{Z}$ , or  $f$  is a trigonometric polynomial of degree  $N$ , and  $\hat{f}(k) \neq 0$  for all  $|k| \leq N$ .
- (ii) For every  $n \geq 3$  (and  $n \leq$  the degree  $N$  in the case of  $f$  being a trigonometric polynomial), we have  $\text{Arg}(\hat{f}(n)) - \text{Arg}(\hat{f}(n-1))$  does not differ from  $\text{Arg}(\hat{f}(2)) - \text{Arg}(\hat{f}(1))$  by an integer multiple of  $\pi$ . Further,  $\text{Arg}(\hat{f}(2)) - 2\text{Arg}(\hat{f}(1))$  is not an integer multiple of  $\pi$ . Here we consider  $\text{Arg}$  to be a number in  $(-\pi, \pi]$ .

We claim that the complement of  $\mathfrak{F}$  in  $\mathcal{E}$ , denoted by  $\mathcal{G} := \mathcal{E} \setminus \mathfrak{F}$ , is a meagre subset of  $\mathcal{E}$ , and for  $f \in \mathfrak{F}$ , we have  $\mathbb{P}^f$  determines  $f$  up to the rotation and flip.

To show that  $\mathcal{G}$  is meagre, we will show that it is a subset of a countable union of nowhere dense sets. Indeed, we can write  $\mathcal{G}$  as:

$$\mathcal{G} \subset \cup_i A_i \cup_{n \geq 3} B_n \cup C$$

where

$$A_i := \{f \in \mathcal{E} : \hat{f}(i) = 0\},$$

$$B_n := \{f \in \mathcal{E} : \text{Either } \hat{f}(k) = 0 \text{ for } k = 1, 2, n, n-1 \text{ or}$$

$$\text{Arg}(\hat{f}(n)) - \text{Arg}(\hat{f}(n-1)) = \text{Arg}(\hat{f}(2)) - \text{Arg}(\hat{f}(1)) + t\pi, t \text{ an integer with } |t| \leq 4\},$$

$$C := \{\hat{f}(1) = 0 \text{ or } \hat{f}(2) = 0 \text{ or } \text{Arg}(\hat{f}(2)) - 2\text{Arg}(\hat{f}(1)) = t\pi, t \text{ an integer with } |t| \leq 3\}$$

It is not hard to see that each  $A_i$  and  $B_n$  are closed sets in  $L^\infty(\mathbb{T})$ , being defined by closed conditions on finitely many co-ordinates of the Fourier expansion (observe that  $|\hat{f}(n)| \leq \|f\|_\infty$  for each  $n$ ), and the same holds true for  $C$ . It is also clear that none of the  $A_i$ -s or  $B_n$ -s or  $C$  contain any  $L^\infty(\mathbb{T})$  ball, showing that they are nowhere dense. All these combine to prove that  $\mathcal{G}$  is a meagre subset of  $L^\infty(\mathbb{T})$ .

Let  $f$  be a function in  $\mathfrak{F}$ . We begin with the Fourier expansion  $f$ :

$$f(x) = \sum_{-\infty}^{\infty} a_j e^{ijx} \quad (18)$$

where  $i$  is the imaginary unit.

We make the following observations about the expansion (18). First,  $f$  is real valued implies

$$a_{-j} = \overline{a_j} \text{ for all } j. \quad (19)$$

Secondly, letting  $f_\xi = f(x + \xi) = \sum_{j=-\infty}^{\infty} a_j^\xi e^{ijx}$  where  $\xi \in (-\pi, \pi]$  and the addition is in  $\mathbb{T}$ , we have

$$a_j^\xi = a_j e^{ij\xi}. \quad (20)$$

Finally, setting  $\tilde{f}(x) = f(-x) = \sum_{j=-\infty}^{\infty} \tilde{a}_j e^{ijx}$ , we have

$$\tilde{a}_j = \overline{a_{-j}}. \quad (21)$$

We want to recover the coefficients  $a_j$  (up to the symmetries (19), (20) and (21)) from the measure  $\mathbb{P}^f$ . We observe that the class of functions  $\mathfrak{F}$  is preserved under these symmetries. In particular, a trigonometric polynomial remains a trigonometric polynomial of the same degree and the same regularity property as demanded in the definition of the class  $\mathfrak{F}$ .

To this end, we begin by observing that  $a_0 = \rho_1$ , and is therefore determined by  $\mathbb{P}^f$ .

Further,

$$\rho_2(0, n) = \begin{vmatrix} a_0 & a_n \\ a_{-n} & a_0 \end{vmatrix} = a_0^2 - |a_n|^2. \quad (22)$$

This implies that  $\mathbb{P}^f$  determines  $|a_n|$  for all  $n$ .

Recall that  $a_1 \neq 0$ . Using the symmetry (20), we choose  $a_1$  to be a positive real number, equal to its absolute value which is determined by  $\mathbb{P}^f$ .

For any integer  $n$ , we have

$$\rho_3(0, p, n) = \begin{vmatrix} a_0 & a_1 & a_n \\ a_{-1} & a_0 & a_{n-1} \\ a_{-n} & a_{-(n-1)} & a_0 \end{vmatrix} \quad (23)$$

Expanding the right hand side along the first row, we have

$$\rho_3(0, 1, n) = a_0 \begin{vmatrix} a_0 & a_{n-1} \\ a_{-(n-1)} & a_0 \end{vmatrix} - a_1 \begin{vmatrix} a_{-1} & a_{n-1} \\ a_{-n} & a_0 \end{vmatrix} + a_n \begin{vmatrix} a_{-1} & a_0 \\ a_{-n} & a_{-(n-1)} \end{vmatrix}.$$

Expanding the  $2 \times 2$  determinants, we can simplify the above equation to

$$\rho_3(0, 1, n) = 2a_1 \Re(a_n \overline{a_{n-1}}) + g(a_0, a_1, |a_{n-1}|, |a_n|), \quad (24)$$

where  $g(w, x, y, z)$  is a polynomial in four complex variables. Since all quantities in (24) except  $\Re(a_n \overline{a_{n-1}})$  are known and  $a_1 > 0$ , we deduce that  $\mathbb{P}^f$  determines  $\Re(a_n \overline{a_{n-1}})$  for all integers  $n$ .

For  $n = 2$ , we have  $a_{n-1} = a_1$ , and this implies that  $\mathbb{P}^f$  in fact determines  $\Re(a_2)$ . Since  $|a_2|$  is also known, this implies that  $\mathbb{P}^f$  determines  $|\Im(a_2)|$ , which is non-zero because of the assumption  $\text{Arg}(\hat{f}(2)) - 2\text{Arg}(\hat{f}(1))$  is not a multiple of  $\pi$ ,  $\text{Arg}(\hat{f}(1))$  being 0 because  $\hat{f}(1)$  is real. Using the symmetry (21), we choose  $a_2$  such that  $\Im(a_2) > 0$ .

We have now spent all the symmetries present in the problem, and our goal is to show that all the other  $a_n$ -s ( $n \geq 0$ ) are determined exactly by  $\mathbb{P}^f$ .  $a_n$  for  $n < 0$  can then be found using the symmetry (19).

To this end, we apply induction. Suppose  $n \geq 3$  we know the values of  $a_k, 0 \leq k \leq n-1$ . Computing  $\rho_3(0, 2, n)$  along similar lines to  $\rho_3(0, 1, n)$  we obtain

$$\rho_3(0, 2, n) = 2\Re(a_n \overline{a_2 a_{n-2}}) + g(a_0, a_2, |a_{n-2}|, |a_n|), \quad (25)$$

where  $g$  is as in (24). Since we already know  $|a_n|$ , we need only to determine  $\text{Arg}(a_n)$ . For any complex number  $z \neq 0$ ,  $|z|$  and  $\Re(z)$  determines  $\text{Arg}(z)$  (considered as a number in  $(-\pi, \pi]$ ) up to sign. Hence, if we know  $\Re(z\bar{z}_1)$  and  $\Re(z\bar{z}_2)$  for two non-zero complex numbers  $z_1$  and  $z_2$  such that  $\text{Arg}(z_1)$  does not differ from  $\text{Arg}(z_2)$  by an integer multiple of  $\pi$ , then this data would be sufficient to determine  $\text{Arg}(z)$ . But this is precisely the situation we have in our hands, with  $z = a_n$ ,  $z_1 = a_{n-1}$  and  $z_2 = a_2 a_{n-2}$ . None of them is 0 by condition (i) defining  $\mathfrak{F}$ , and the condition on the difference of arguments of  $z_1$  and  $z_2$  follows from condition (ii) defining  $\mathfrak{F}$ . This enables us to determine  $\text{Arg}(a_n)$ , and hence  $a_n$ .

This completes the proof. ■

**Remark 5.1.** *The argument used in Proposition 1.6 can also be used to recover the function  $f$  if all its Fourier coefficients are real. E.g., if  $f$  is the indicator function of an interval  $A$  in  $(-\pi, \pi]$ , then we can “rotate”  $f$  (symmetry 20) so that 0 is in the centre of the interval  $A$ . Then all the Fourier coefficients of  $f$  are real, and we can identify the interval  $A$ . This argument will also work for any set  $A$  which has a point of reflectional symmetry when looked upon as a subset of  $\mathbb{T}$ .*

*Proof of Proposition 1.7.* Recall that the harmonic mean of a function  $f : \mathbb{T}^d \rightarrow \mathbb{R}$  is defined as

$$\text{HM}(f) = \left( \int_{\mathbb{T}^d} \frac{d\lambda_d(x)}{f(x)} \right)^{-1}.$$

The fact that we know the distribution  $\mathbb{P}^f$  implies that we know the distribution  $\mathbb{P}^{tf}$  for any  $0 < t < 1$ , e.g. by performing an independent site percolation with survival probability  $t$  on  $\mathbb{P}^f$ . By taking complement, this implies that we know the distribution  $\mathbb{P}^{1-tf}$ . But this enables us to recover the harmonic mean  $\text{HM}(1 - tf)$  of the function  $1 - tf$  by the formula

$$\text{HM}(1 - tf) = \text{Sup}\{p \in [0, 1] : \mu_p \preceq_f \mathbb{P}^{1-tf}\}.$$

Here  $\mu_p$  is the standard site percolation on  $\mathbb{Z}$  with survival probability  $p$ , and  $\mu_p \preceq_f \mathbb{P}^{1-tf}$  means that  $\mathbb{P}^{1-tf}$  is uniformly insertion tolerant at level  $p$ , that is,  $\mathbb{P}^{1-tf}[\omega(0) = 1 | \omega_{\text{out}}] \geq p$  a.s. in  $\omega_{\text{out}}$ . For details, we refer to Definition 5.15 and Theorem 5.16 in [LySt03]. But

$$\text{HM}(1 - tf)^{-1} = \int_{\mathbb{T}^d} \frac{d\lambda_d(x)}{1 - tf(x)}.$$

By expanding the integral on the right as a power series in  $t$ , we can recover all moments  $\int_{\mathbb{T}^d} f^k(x) d\lambda_d(x)$  of  $f$ . But we have

$$\int_{\mathbb{T}^d} f^k(x) d\lambda_d(x) = \int_0^1 \xi^{k-1} \nu_f(\xi) d\mathcal{L}(\xi),$$

where  $\mathcal{L}$  is the Lebesgue measure on  $\mathbb{R}$ , and  $\nu_f$  is the value distribution of  $f$ , given by

$$\nu_f(\xi) = \lambda_d(\{x \in \mathbb{T}^d : f(x) \geq \xi\}).$$

This enables us to recover the value distribution of  $f$ . ■

**Acknowledgements.** We would like to thank Russell Lyons, Vadim Gorin and Yuval Peres for stimulating discussions.

## References

- [AGZ09] G. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press, 2009.
- [CL97] G. Chistyakov, Y. Lyubarskii, *Random perturbations of exponential Riesz bases in  $L^2(-\pi, \pi)$* , Ann. de l'institute Fourier, tome 47, no. 1, 201-255 (1997).
- [CLP01] G. Chistyakov, Yu. Lyubarskii, L. Pastur, *On the completeness of random exponentials in the Bargmann-Fock space*, Journal of Mathematical Physics, Vol. 42 No. 8 (2001).
- [DV97] D.J. Daley, D. Vere Jones, *An Introduction to the Theory of Point Processes (Vols. I & II)*, Springer, 1997.
- [GP12] S. Ghosh, Y. Peres, *Rigidity and Tolerance in point processes: Gaussian zeroes and Ginibre eigenvalues*, in preparation.
- [HKPV10] J.B. Hough, M. Krishnapur, Y. Peres, B. Virag, *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, A.M.S., 2010.
- [HS10] A.E. Holroyd, T. Soo, *Insertion and deletion tolerance of point processes*, arXiv:1007.3538v2, 2010.
- [Ly03] R. Lyons, *Determinantal Probability Measures*, Publ. Math. Inst. Hautes Etudes Sci., 98 (2003), 167-212.
- [LySt03] R. Lyons, J. Steif, *Stationary determinantal processes: phase multiplicity, Bernoullicity, entropy, and domination*, Duke Math. J. 120, no. 3, 515-575 (2003).
- [Mo03] B. Morris, *The Components of the Wired Spanning Forest are Recurrent*, Prob. Theor. Rel. Fields, 125, pp. 259-265, 2003.
- [Re77] R. Redheffer, *Completeness of Sets of Complex Exponentials*, Adv. Math., 24, 1-62 (1977).

- [ReYo83] R. Redheffer, R. Young, *Completeness and basis properties of complex exponentials*, Trans. Amer. Math. Soc., Vol. 277 No. 1 (1983).
- [Se92] K. Seip, *Density theorems for sampling and interpolation in the Bargmann-Fock space*, Bull. Amer. Math. Soc., Vol. 26 No. 2, 322-328 (1992).
- [Sos00] A. Soshnikov, *Russian Math. Surveys* 55:5 923-975 (*Uspekhi Mat. Nauk* 55:5 1071-160), 2000.
- [SU97] K. Seip, A. Ulanopvskii, *The Beurling-Malliavin density of a random sequence*. Proc. Amer. Math. Soc., Vol. 125, No. 6, 1745-1749 (1997).
- [ShTa03] T. Shirai, Y. Takahashi, *Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes*, Journal of Func. Anal., 205, 414-463, (2003).